The Regulatory Pathway for Rare Diseases
Lessons Learned from Examples of Clinical Study Designs for Small Populations

Robert J. Temple, M.D.

Deputy Center Director for Clinical Science

Center for Drug Evaluation and Research

U.S. Food and Drug Administration

Symposium: Best Practices in Clinical
Study Design for Rare Diseases
April 29, 2013

Law and Regulations

The basis for approval under law for all drugs, including orphan drugs, is "substantial evidence that the drug will have the effect it . . . is represented to have [under labeled conditions]. Substantial evidence means evidence consisting of adequate and well-controlled clinical investigations [that allow a conclusion] that the drug will have the effect it. . . is represented to have." This generally means two studies but under FDAMA (1997) substantial evidence can be based on a single A&WC study with "confirmatory evidence" (not further defined).

Regulations at 314.50 describe 5 possible controls: placebo, no-treatment, active, dose-response, and historical and ICH E-10 [Choice of Control Group and Related Issues in Clinical Trials] describes when a historical control can be credible.

Law and Regulations

Although regulatory requirements are the same, Subpart E of the IND regulations (part 312.80-88) describes special considerations for Drugs Intended to Treat Life-Threatening and Severely-Debilitating Illnesses, including "the broadest flexibility in applying the statutory standards."

A critical area of potential flexibility is reliance on a single study. In a 1998 guidance "Providing Clinical Evidence of Effectiveness," FDA described the kind of confirmatory evidence that could allow reliance on a single study; in many cases this evidence would be other studies in related diseases but it could also be a well-documented mechanistic/pharmacologic effect that was clearly related to the etiology and mechanism of the disease, e.g., replacement of a missing enzyme or other protein with well-defined activity. Of course, the Accelerated Approval mechanism (subpart H of section 314) allows initial approval based on an effect on a surrogate endpoint such as replacement of an enzyme function.

Past Activity

FDA has used the mechanisms suggested under Subpart E (312.80) and Subpart H (314.510) fairly often as a basis for approval for orphan drugs. The "score" is described in a paper by Frank Sasinowski of NORD published in 2011; setting forth the "Quantum of Effectiveness Evidence in FDA's Approval of Orphan Drugs." Apart from demonstrating flexibility, it suggests the pathways that FDA has used and that therefore can be considered, several of which I will expand on

Past History

The NORD report makes interesting reading, but the cases I would emphasize are these:

1. Historically controlled studies, including both comparisons of outcomes on drug and an explicit non-treated group over a similar time period, and baseline controlled studies in which the treated group is compared with the "natural history" of the disease, but not a specific group of patients. This is common in oncology, where tumor response rates are often a basis for accelerated approval (or even full approval) because it is clear that tumors do not shrink in the absence of treatment. It is clear that a major issue in these cases is the quality of the data defining what actually occurs in the absence of therapy.

2. Very small randomized studies that are nonetheless successful in showing effects, sometimes the only study, and sometimes the second study, supporting effectiveness.

Historical Controls
Natural History of the Disease

There are two critical reasons for trying to define as well as possible the natural history of an orphan disease and the variability of that history.

1. If the natural history is very well defined and the drug effect is large, a single-arm, historically controlled study (baseline controlled or with an explicit historical control group) can be, and sometimes has been, the basis for approval.

BUT, you really must know the history, as a long-past example illustrates. It suggests that one must always be concerned about even natural histories that seem clearly defined.

Historical Controls
Fulminant Hepatitis B

In a letter to the NEJM in 1971, Gocke described 9 cases of acute fulminant hepatitis B, all fatal despite exchange transfusion, steroids. Then they gave 8 patients with hepatic coma or pre-coma anti-Australia antigen serum with 5/8 survival.

Gocke thought, maybe, they were done, and was concerned about the ethics of doing a controlled trial, but in the end he was unsure about whether current patients and Rx were all the same so they did an NHLBI-sponsored RCT of hyperimmune globulin vs normal serum in 30 centers with 63 patients, 53 of whom were analyzed (10 did not have hepatitis B antigen or had no specimens).

Survival was 9/28 (32%) on placebo and 7/25 (28%) on hepatitis B immune globulin, surely a surprise in view of the historical experience [Acute Hepatic Failure Study Group. Failure of specific immunotherapy in fulminant type B hepatitis. Amer. Int Med (1977); 86: 272-277.]

Historical Controls

Critical Reference -

      Sacks, Chalmers, Smith

      Am J. Medicine (1982); 72:233-240.

Comparison of RCTs and HCTs for same disease

<u>Always</u>

1. RCT less favorable than HCT
2. Reason was that the historical control was worse than the randomized control (selection bias)
3. Not possible to "adjust" the difference

Many examples of misleading HCTs; great care in relying on one

Table I - Conclusions of RCTs and HCTs on Six Therapeutic Questions

| Question Studied	RCT Effective	RCT Ineffective	HCT All Trials Effective	HCT All Trials Ineffective	HCT Matched or Adjusted for Prognostic Factors Effective	HCT Matched or Adjusted for Prognostic Factors Ineffective
Cirrhosis with Varices	6	14	12	6	2	1
Coronary Artery Surgery	1	7	16	5	9	1
Anticoagulants for Acute Myocardial Infarction	1	9	5	1	3	1
5-FU Adjuvant for Colon Cancer	0	5	2	0	2	0
BCG Adjuvant for Melanoma	2	2	4	0	4	0
DES for Habitual Abortion	0	3	5	0	1	0
TOTALS	10	40	44	12	21	3
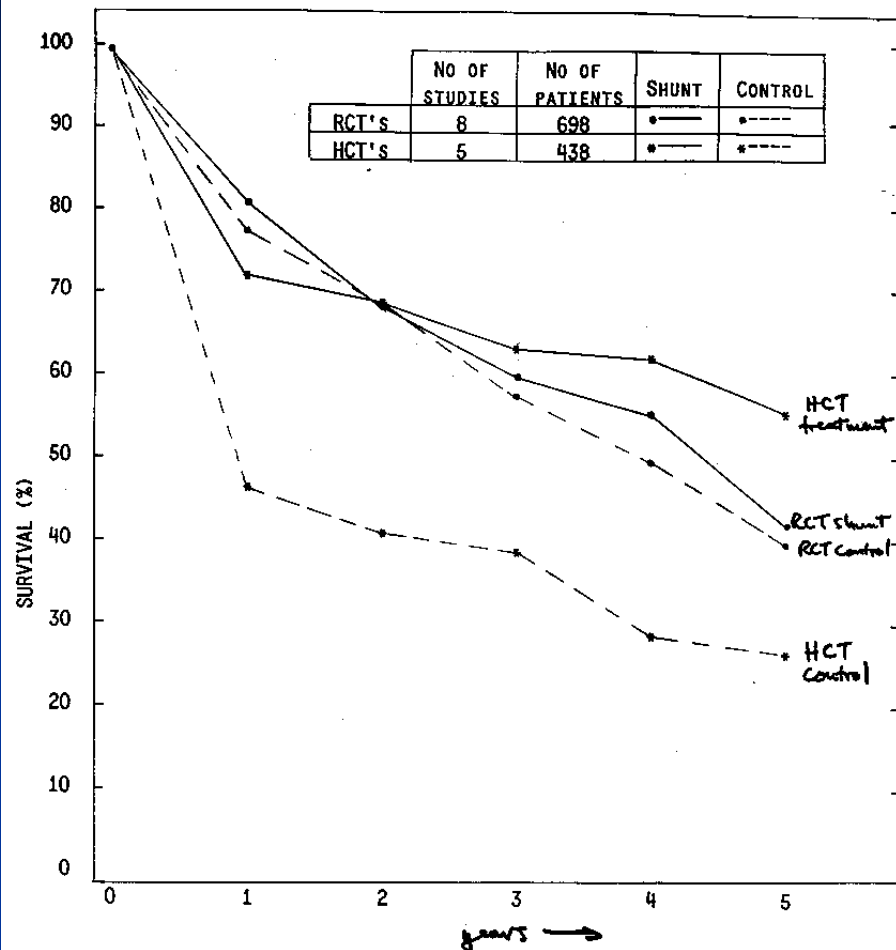
FIGURE 1. SURVIVAL OF TREATED AND CONTROL GROUPS IN CLINICAL TRIALS OF SHUNT SURGERY FOR CIRRHOSIS WITH ESOPHAGEAL VARICES.

Historical Controls

So care in use of historical controls is critical. It greatly helps when mechanism of disease is clear and drug is designed to reverse/correct the mechanism.

2. The second major reason to understand the natural history is that it will help greatly in designing arandomized trial if one is needed.

A Concurrently Controlled Clinical Trial Is Needed

In many cases, even for orphan diseases, the natural history is not "fixed" enough or well known enough and, as my examples illustrate, these designs have problems, and a randomized trial will be needed (one or 2, separate question). It is therefore critical to think about how to do that most efficiently. There are 2 specific designs to consider:

- Enrichment designs
- Crossover and N of 1 designs

Efficiencies in Concurrently Controlled Trials

But first, a pitch:

As I've noted, and as NORD documents, we have relied on historical controls and they can indeed be persuasive, but there will usually be a "kernel" of doubt, always a concern for a drug developer, so it is critical to ask: can this be avoided?

I recall, 20-30 years ago, hearing Tom Chalmers urge people, especially where bad diseases were involved, to "randomize the first patient," because later, if there were hints of effect, there will be growing reluctance to do so. I just read the invitation to the Cochrane Colloquium in Canada in Sept and their plans for the Thomas C. Chalmers Award, which says

> He is perhaps best known for the notion "randomize the first patient," reflecting the belief that it is more ethical to randomize than to treat in the absence of good evidence.

Amen.

Efficiencies

Let me also note a recent publication by Korn, McShane, and Friedlin of the NCI [Statistical challenges in the evaluation of treatments for small patient population. Science Translational Medicine 2013; 5:1-14] that discusses the full range of design considerations for studies of diseases with small patient populations.

Enrichment:

> In various ways enriched studies seek to test therapies in patients with a high likelihood of having an endpoint (prognostic enrichment) or of having a response (predictive enrichment), each of which allows for a smaller study. One specific design is the randomized withdrawal study, where patients doing well on a treatment are randomized to continued treatment or placebo, a kind of predictive enrichment.
> The first study of this type we saw was the basis for approval for nifedipine for vasospastic angina.

Nifedipine Randomized Withdrawal
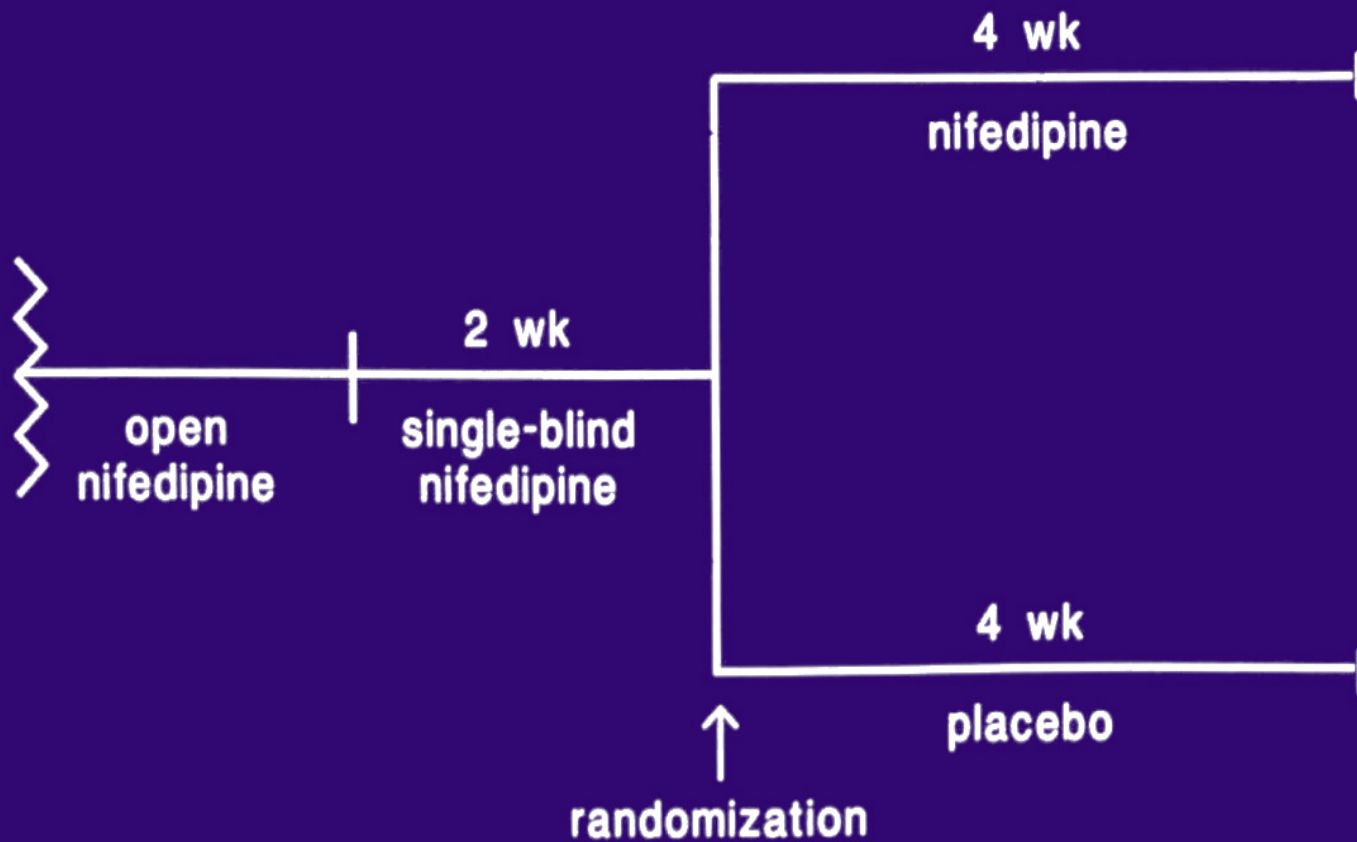
	Nifedipine	Placebo
n	13	15
Early withdrawal	0	5*
Early withdrawal plus AMI	0	6*
Investigator's judgment of success	11	2*
Median angina/week	0	3.4*
Mean angina/week	0.7	18.4*

Change from baseline in attacks/week

	Nifedipine	Placebo
better (≤ 1)	0	0
same (± 1)	11	6
worse (≥ 1)	2	9

*$p < 0.05$, one sided

Enrichment

In two cases, tetrabenazine for treatment of choreiform movements in Huntington's Chorea and sodium oxybate for cataplexy, a second supportive study used this design.

Enrichment

Patients on treatment with sodium oxybate for cataplexy with narcolepsy for 7-44 months randomized to continued treatment of placebo

|	median attacks/2 weeks	
	Baseline	Change in Rate
Placebo (29)	4.0	+21.0
sodium oxybate (26)	1.9	0

p<0.001

Clearly demonstrated persisting long-term effect and provided a confirmatory trial in practically no time.

Enrichment

It is very common to have, during development of an orphan drug, patients receiving drug in an open-label study, making the randomized withdrawal study possible (not of course, if patients would be harmed).
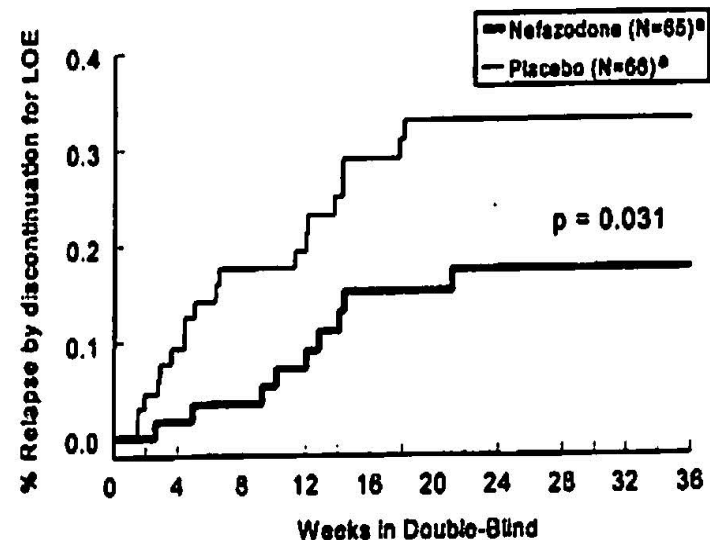
Enrichment

As a further illustration of the power of this design, consider depression. Studies of acute depression fail about 50% of the time. Studies of known responders to establish maintenance effects (reduced recurrence rates), in contrast, show large, regular effects.
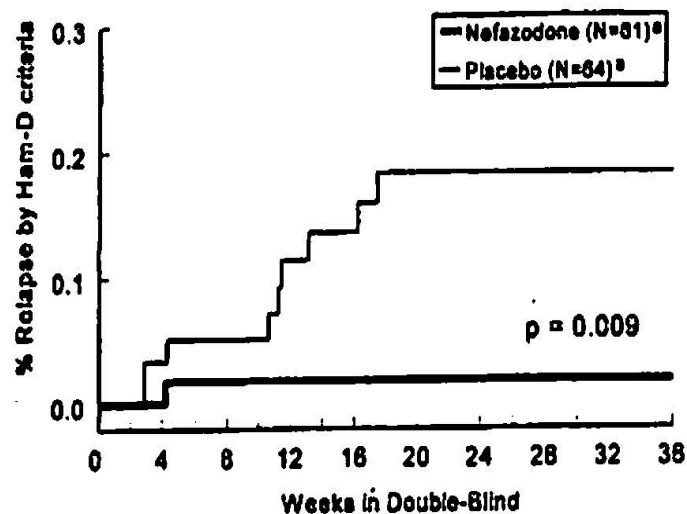
* 38 nefazodone and 34 placebo treated patients finished that trial

* 38 nefazodone and 34 placebo treated patients finished the trial

Figure 1. Kaplan–Meier plots of time to relapse based on criterion of two Hamilton Depression Scale (HAMDS) observations ≥ 18 that were 1 week apart.

Figure 2. Kaplan–Meier plots of time to relapse based on criterion of discontinuation for lack of efficacy (LOE).

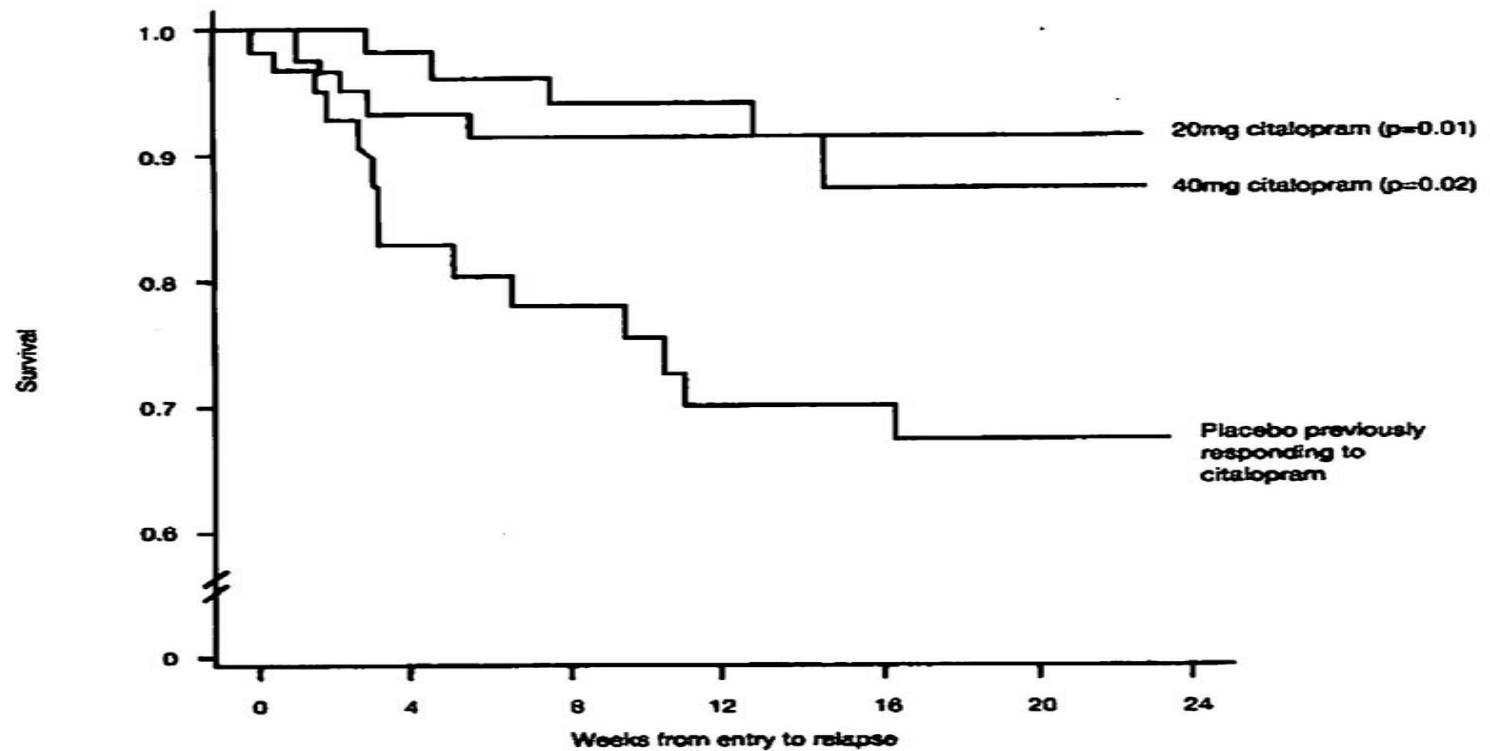FIG. 2. Survival analysis of relapse rates (Kaplan-Meier).

20mg citalopram (p=0.01)

40mg citalopram (p=0.02)

Placebo previously responding to citalopram

Modified Study Designs

1. Cross-over Studies

For a persistent disease, where the drug modifies symptoms or the underlying disease in a reversible way, a randomized cross-over study should decrease the needed sample size by about a factor of two, as each subject serves as own control.

All subjects are exposed to both treatments.

Should minimize inter-subject variability

Would appear attractive in such conditions
- epilepsy
- chronic pain
- many metabolic abnormalities
- diabetic control

23

Modified Study Designs

2. N of 1 designs

Really a kind of x-over. A classic study was Gelfand, et al in 1976 NEJM, a study of danazol in HAE.

Nine patients (with attacks of ≥ 1 per month) were assigned to a random sequence of drug or placebo, to be taken for 1 month, but treatment was stopped if there was an attack, and patient moved to the next treatment.

Total of 46 or 47 course of drug or placebo

    1/46 danazol courses had an attack
    44/47 placebo courses had an attack

The p-value was described as < 0.01, but that was the per patient result. Pooled would be far smaller.

Note 2 things
A tiny number of patients had MANY treatments
It worked out because effect was large

Conclusion

Where a clinical trial is needed, it is far safer to use a concurrently controlled trial than to rely on a historical control unless the effect is very large. If the effect is large, stopping rules can limit the duration and study size so that little time will be wasted. And if there is uncertainty about uniformity of patients with the disease, the control protects against surprises and uncertainty. There seems little reason not to make the first patient trial an RCT, with rare exception (refractory cancer, where response rate is wholly assessable without a control).